CHEM 371: Lecture 2

● After the long introductory remarks of Lecture 1, we're finally in a position to say something definite about the kinds of questions that time-dependent statistical mechanics will try and address. The first thing to recognize is that time-dependent statistical mechanics, like its time-independent counterpart seeks to understand the results of experiments or simulations in microscopic/molecular terms. (In other words, it's looking to explain phenomena theoretically.) So we have to first ask what kinds of experiments fall within the purview of the subject.

One kind of experiment is the kind we introduced at the outset to illustrate what we meant by microstate, macrostate and equilibrium. We'll now revisit that example but in somewhat more general terms. So this is the experiment (adapted from notes provided by Hans Anderson, Stanford University):

– Take some system (solid, liquid, gas), prepare it in some specific way (e.g., by fixing its temperature, or placing it in a magnetic field), and then let it attain equilibrium (where its macroscopic variables stop changing.)
– At this point, set the time to 0, and record the value of a dynamical variable *A*. From what we've said about the duration of experimental measurements, this value is effectively a time average of *A*, but we won't worry about what this value actually corresponds to; we'll simply regard it as an experimental data point taken at time 0, and refer to it as $A(0)$.
– Repeat these steps several times. (To check reproducibility.) Average the set of $A(0)$ values so obtained. The average is the experimental estimate of *A*, and we'll call it $\overline{A}(0)_{\text{exp}}$. (Deviations from this average in individual measurements can be reported as error bars.) We will now assert, as we did earlier, that the same estimate is obtained from an ensemble average of *A*, which can be determined using statistical mechanics. Specifically,

$$\begin{aligned}\overline{A}(0)_{\text{exp}} &= \langle A(0)\rangle \\ &= \int d\Gamma(0) A(\Gamma(0)) P(\Gamma(0)) \\ &\equiv \int d\Gamma A(\Gamma) P(\Gamma) \quad (d\Gamma = dq_1 dp_1 \cdots dq_N dp_N)\end{aligned}$$

where $P(\Gamma)$ is the probability that at $t = 0$ the system is in the microstate Γ (the probability, in other words that the position and momentum of particle 1 are $\mathbf{q}_1$ and $\mathbf{p}_1$, those of particle 2 are $\mathbf{q}_2$ and $\mathbf{p}_2$, and so on for the all the other particles in the system.) If the system under consideration happens to be closed (as in the example we discussed earlier of a gas in a sealed container), every microstate is assumed to be equally likely (an assumption referred to as the assumption of equal à priori probabilities). In such a system $P(\Gamma) = 1/\Omega$, where Ω is the total number of microstates available to the system. If the system is not closed, however (if it is in thermal contact with its surroundings, for instance, so that its temperature *T* rather than its energy *U* is a constant), the microstates available to the system are now no longer equally probable; some microstates will be more likely than others. The precise form of $P(\Gamma)$ is fixed by the macroscopic constraints on the system

(such as constant T or P, etc.), and can be determined from equilibrium statistical mechanics.

In another kind of experiment, we'll do the following:
– Prepare the system in some specific way.
– Allow the system to equilibrate.
– When the system has equilibrated, set the time to 0.
– Allow the system to evolve for a time t.
– Record the value of some property A at this time. Denote this value $A(t)$.
– Repeat these steps several times, and average the different values of $A(t)$ to obtain the experimental estimate of this quantity, which can be denoted $\overline{A}(t)_{\text{exp}}$. The same result is can be obtained from a statistical mechanical calculation according to the prescription below:

$$\begin{aligned}\overline{A}(t)_{\text{exp}} &= \langle A(t) \rangle \\ &= \int d\Gamma(t) A(\Gamma(t)) P(\Gamma(t))\end{aligned}$$

We'll prove later that since the system is in equilibrium, $P(\Gamma(t)) = P(\Gamma(0))$. (This seems intuitively reasonable.) We'll also prove that $d\Gamma(t) = d\Gamma(0)$. Therefore,

$$\begin{aligned}\langle A(t) \rangle &= \int d\Gamma(0) A(t) P(\Gamma(0)) \\ &= \int d\Gamma A(t) P(\Gamma)\end{aligned}$$

Now what we mean by $A(t)$ is the value of A at t given that it had started at $t = 0$, where the system was in the microstate $\Gamma(0) = \Gamma$. This circumstance can be made explicit by writing $A(t)$ as $A(t;\Gamma)$. Hence

$$\langle A(t) \rangle = \int d\Gamma A(t;\Gamma) P(\Gamma)$$

Because the system is in equilibrium, it should be the case that the average of A determined at time t is the same as its average determined at any other time. Referring to the above equation, this means, in effect, that

$$\langle A(t) \rangle = \langle A(0) \rangle$$

Notice that what is being asserted is not that $A(t;\Gamma) = A(0)$ but that their *averages* over the same equilibrium distribution of microstates are equal. Again, this seems intuitively reasonable. As we pointed out earlier, a measurement of some property A (and by measurement I mean effectively a time average) that is carried out today should produce the same result if it were carried out tomorrow, provided the system is in equilibrium, and it remains visibly unchanged between now and the next day.

This general circumstance, that $\langle A(t)\rangle = \langle A(0)\rangle$ for equilibrium averages is called the property of *stationarity*.

In both the above experiments, the method used to determine an ensemble average coincides exactly with the general formalism of equilibrium statistical mechanics. So the $P(\Gamma)$'s are nothing but the various probability distributions we've encountered before, such as the canonical distribution or the grand canonical distribution, and so on.

We'll conduct one other type of experiment:
– Prepare the system in some specific way.
– Allow the system to equilibrate.
– When the system has equilibrated, set the time to 0.
– Allow the system to evolve for a time t and measure A at this time; denote its value is $A(t)$.
– Allow the system to evolve to time t' and measure another property, say B, at this time, denoting it $B(t')$.
– Multiply the two measured values together.
– Repeat these steps a number of times and average the data. The result is $\overline{A(t)B(t')}_{\text{exp}}$, and as before, we'll assert that it can be obtained from statistical mechanics via an ensemble average:

$$\begin{aligned}\overline{A(t)B(t')}_{\text{exp}} &= \langle A(t)B(t')\rangle \\ &= \int d\Gamma A(t;\Gamma)B(t';\Gamma)P(\Gamma)\end{aligned}$$

And again, because these measurement are made under equilibrium conditions, it doesn't really matter when you start the clock ticking – what matters is how much later the measurement of B is in relation to A, i.e., only the difference in time between these measurements is important. And so

$$\begin{aligned}\langle A(t)B(t')\rangle &= \langle A(0)B(t'-t)\rangle \\ &= \int d\Gamma A(0)B(t'-t;\Gamma)P(\Gamma)\end{aligned}$$

An equilibrium average of the kind

$$C_{AB}(t'-t) = \langle A(t)B(t')\rangle$$

is called an equilibrium *time correlation function* (TCF), and it is this quantity and others like it that will be the focus of one part of time-dependent statistical mechanics, the part we'll study this quarter. (A later part – to be taken up in the next quarter – will consider time-dependent systems whose dynamics are not centered around equilibrium time correlation functions.) We'll see that TCFs play somewhat the same kind of role in time-

dependent statistical mechanics that partition functions play in equilibrium statistical mechanics, and we'll find that even though we're dealing with systems in equilibrium, TCFs are descriptive of a large number of experimentally relevant time-dependent situations. If in the above correlation function, the property *B* happens to be *A* itself, the result, $C_{AA}(t'-t)$, is referred to as an *equilibrium time autocorrelation function*.

● Time Evolution of Microstates

It should be clear from the foregoing discussion that to calculate an equilibrium time correlation function, we'll need the following ingredients:

– An expression for the distribution function $P(\Gamma)$, which tells you the likelihood (probability) that a system of *N* particles under some set of fixed control parameters (like constant *T* or *P*, etc.) is found in some definite microstate $\Gamma = \{q_1, p_1, \ldots, q_N, p_N\}$,
– The value of some property *A* of the system at the time *t* when the system had started out at $t = 0$ in the microstate Γ, i.e., we need an expression for the function $A(t;\Gamma)$. If the TCF we're calculating is $\langle A(t)B(t')\rangle$, we'll need an expression for $B(t';\Gamma)$ too.

We'll look first at the problem of calculating the value of a dynamical variable *A* at some time *t*. Now *A* is a function of Γ, i.e., it depends on the positions and momenta of the particles that make up the system. We've assumed these particles to be classical, so their motion is governed, in general, by Newton's laws, viz.,

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \mathbf{F}_i(\mathbf{r}_1, \ldots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \ldots, \dot{\mathbf{r}}_N, t), \qquad i = 1,2, \ldots \ldots, N \tag{1}$$

where m_i is the mass of the *i*th particle, $\mathbf{r}_i$ is its position (in Cartesian coordinates), and $\mathbf{F}_i$ is the net force acting on it. Given the initial positions and velocities of these particles and the forces on them, these equations can be solved in principle for the positions and velocities at any future time. So if it's known how *A* depends on the $\mathbf{r}_i$, in principle, its value at *t*, viz., $A(t;\Gamma)$ can be determined.

But Newton's laws aren't always easy to apply. For one thing, the position $\mathbf{r}_i$ must be a Cartesian vector in Euclidean space. It can't be some more general variable, like an angle, which might be a more appropriate or natural variable to use in certain circumstances. (If we did use such variables, their evolution equations wouldn't necessarily take the familiar $F = ma$ form of Newton's laws.) A second shortcoming of the Newtonian approach is that it can be difficult to correctly identify all the forces that act on a particle.

Fortunately, there are other formulations of classical mechanics besides Newton's that don't suffer from these limitations. One of them is Lagrangian mechanics. It's based on what's come to be known as the principle of least action, which states that the motion

of a system of particles is such that it minimizes a quantity called the *action*, *S*, which is defined as follows:

$$S = \int_0^t dt' L(\mathbf{q}_1, \ldots, \mathbf{q}_N, \dot{\mathbf{q}}_1, \ldots, \dot{\mathbf{q}}_N, t) \tag{2}$$

where *L* is the difference between the kinetic energy *K* and the potential energy *V* of the system, i.e.,

$$L = K - V\,, \tag{3}$$

and is referred to as the *Lagrangian*. The coordinates and velocities in *L* need not be Cartesian coordinates in Eucidean space – they can be any convenient parameters that uniquely specify the configuration of the system. And they are represented as vectors to indicate that they're not restricted to $1d$. Most of the time we'll be interested in $3d$ systems.

The requirement that *S* be a minimum under the least action principle will impose conditions on these generalized positions and velocities that will determine how they evolve in time. But how do we minimize a function like *S* that involves an integral? We know that for an ordinary function like

$$x = x(t)$$

the necessary condition for it to have a minimum is that its first derivative be 0, i.e., $\dot{x}(t) = 0$. The solution of this equation is some particular value of *t*, say, $t_m = a$. But *S* is not a simple function – it's an integral over another function, which makes it something we call a *functional*. Minimizing or maximizing a functional, or finding the condition that minimizes or maximizes it means finding a *function* that when put back in the expression for it leads to a value that is an extremum (i.e., a maximum or a minimum.) How do we find this function, or what are the conditions that determine it?

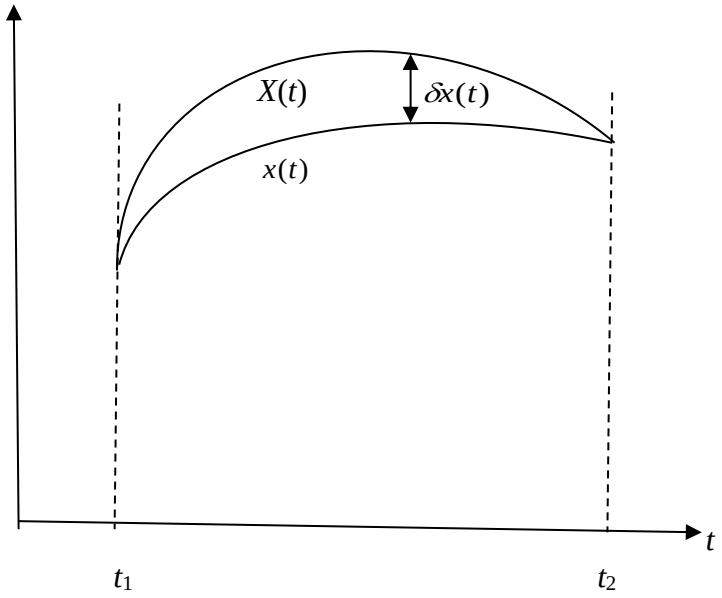
As an illustration of the general procedure, let's consider this representative problem in $1d$. We have some function f that depends on these variables: the time *t*, a generalized coordinate *x* that itself depends on *t*, and a generalized velocity *v*, which is just the time derivative of *x*, and which also depends on *t*. In other words,

$$f = f(x, \dot{x}, t)$$

Given this function, let's construct the following integral:

$$I = \int_{t_1}^{t_2} dt\, f(x(t), \dot{x}(t), t) \tag{4}$$

and then ask how x would have to depend on t in order that I have a *stationary* value when the limits t_1 and t_2 are fixed. What we mean by stationary will become clear shortly. But what should immediately be clear is that the value of this integral depends on *how* x gets from t_1 to t_2 as t is continuously varied between these limits, i.e., it depends on x's path. The largest or smallest such value of I is obtained for one special path, which we'll denote $x(t)$ and refer to as the stationary path. Now consider some other path that x could have taken – call it $X(t)$ – and assume that it's not all that different from the stationary path along the entire length of the given time interval. What that means is schematized in the figure below:



Here $\delta x(t) \equiv X(t) - x(t)$, that is, the difference between the "adjacent" path and the stationary path for all t in the interval between t_1 and t_2, with both paths starting and ending at the indicated points. We'll assume that $\delta x(t)$ is small, in fact, infinitesimal. Given these two neighbouring paths, we'll introduce another infinitesimal quantity, viz.,

$$\delta f \equiv f(X, \dot{X}, t) - f(x, \dot{x}, t) \tag{5}$$

In the above expression as well as in the expression for $\delta x(t)$, the symbol δ stands for "variation", and it represents the change in the value of a function when the function is displaced from its original path for some fixed value of t.

Consider the variation of the derivative of x, $dx/dt \equiv \dot{x}$. By definition

$$\delta \frac{dx}{dt} = \frac{dX}{dt} - \frac{dx}{dt} = \frac{d}{dt}(X - x) = \frac{d}{dt}\delta x \tag{6}$$

So in a formal sense, the operations of variation and differentiation "commute". Now consider the variation of *f* itself, as defined by Eq. (5):

$$\delta f = f(x + \delta x, \dot{x} + \delta\dot{x}, t) - f(x, \dot{x}, t)$$

If we Taylor expand this expression around $\delta x \to 0$ and $\delta\dot{x} \to 0$, we get

$$\begin{aligned} \delta f &= f(x, \dot{x}, t) + + \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial \dot{x}}\delta\dot{x} - f(x, \dot{x}, t) \\ &= \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial \dot{x}}\delta\dot{x} \end{aligned} \tag{7}$$

which shows that rules for calculating variations are formally the same as those for calculating differentials.

To return to the question of what condition determines whether the integral *I* in Eq. (4) is at a maximum or a minimum (i.e., whether it is *stationary*), it is clearly that *I* evaluated along the extremum path *x*(*t*) must be yield the same value as *I* evaluated along the infinitesimally displaced neighbouring path $x + \delta x$. In other words,

$$\int_{t_1}^{t_2} dt\, f(x + \delta x, \dot{x} + \delta\dot{x}, t) = \int_{t_1}^{t_2} dt\, f(x, \dot{x}, t) \tag{8}$$

That is,

$$\int_{t_1}^{t_2} dt\, \delta f(x, \dot{x}, t) = 0 \tag{9}$$

This is the counterpart of the condition that determines stationarity in ordinary calculus, viz., $dx = 0$. If you substitute (6) and (7) into (9), the result is

$$\int_{t_1}^{t_2} dt \left(\frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial \dot{x}}\frac{d}{dt}\delta x \right) = 0 \tag{10}$$

and if you then integrate the second term in (10) by parts, you get

$$\int_{t_1}^{t_2} dt \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial \dot{x}}\delta x \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \delta x \frac{d}{dt}\frac{\partial f}{\partial \dot{x}} = 0 \tag{11}$$

The surface term vanishes at the upper and lower limits because there the extremal path and the displaced path both coincide (by construction, see the figure.) So Eq. (11) reduces to

$$\int_{t_1}^{t_2} dt \left(\frac{\partial f}{\partial x} - \frac{d}{dt} \frac{\partial f}{\partial \dot{x}} \right) \delta x = 0 \tag{12}$$

Now, one way this equation could be satisfied is if the product of the deviation δx and the term in parentheses were positive in some interval of time and negative in some others, such that in the entire interval between t_1 and t_2 the negative and positive contributions exactly cancelled one another. But δx is arbitrary, so it could be chosen to have values in certain time intervals that ensured that its product with the parenthetical term didn't produce counterbalancing positive and negative terms, making the integral in (12) non-zero. Therefore, the only way the integral is guaranteed to vanish identically for all δx is if the following condition holds

$$\frac{\partial f}{\partial x} - \frac{d}{dt} \frac{\partial f}{\partial \dot{x}} = 0 \tag{13}$$

The solution of this equation provides the functional form of the extremal path. Eq. (13) is known as the *Euler equation*.

● Several dependent variables

Suppose the integral to extremize is a functional of several dependent variables, as in

$$I = \int_{t_1}^{t_2} dt \, f(x, y, z, \ldots, \dot{x}, \dot{y}, \dot{z}, \ldots\ldots, t) \tag{14}$$

What condition now determines the structure of the extremal paths $x(t)$, $y(t)$, $z(t)$, etc.? If the same steps are followed as before, that is, if *x*, *y*, *z* etc., representing the extremal path, are displaced by small amounts δx, δy, δz, etc., it's easy to show that the stationarity condition is given by

$$\int_{t_1}^{t_2} dt \, \delta f(x, y, z, \ldots, \dot{x}, \dot{y}, \dot{z}, \ldots, t) = 0 \tag{15}$$

where

$$\delta f = \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial z} \delta z + \cdots + \frac{\partial f}{\partial \dot{x}} \delta \dot{x} + \frac{\partial f}{\partial \dot{y}} \delta \dot{y} + \frac{\partial f}{\partial \dot{z}} \delta \dot{z} + \cdots \tag{16}$$

If you substitute (16) into (15), use partial integration again, and eliminate the surface terms, what you're left with is

$$\int_{t_1}^{t_2} dt \left[\left(\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial \dot{x}} \right) \delta x + \left(\frac{\partial f}{\partial y} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}} \right) \delta y + \left(\frac{\partial f}{\partial z} - \frac{d}{dt}\frac{\partial f}{\partial \dot{z}} \right) \delta z + \cdots \right] = 0 \qquad (17)$$

Again, the only way this relation can be satisfied for arbitrary δx, δy, δz, etc., is if each of the parenthetical terms vanishes independently

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial \dot{x}} = 0, \qquad \frac{\partial f}{\partial y} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}} = 0, \qquad \frac{\partial f}{\partial z} - \frac{d}{dt}\frac{\partial f}{\partial \dot{z}} = 0, \ \cdots \qquad (18)$$